数理解析研究所講究録
第 1832 巻 2013 年 59-72

Bounded arithmetic in free logic

独立行政法人 産業技術総合研究所 山形頼之
Yoriyuki Yamagata
National Institute of Advanced Science and Technology (AIST)
yoriyuki.yamagata@aist.go.jp

Abstract

One of the central open questions in bounded arithmetic is whether Buss' hierarchy of theories of bounded arithmetic collapses or not. In this resume, we summarize the author's recent attempt to this problem. We reformulate Buss' theories using *free logic* and conjecture that such theories are easier to handle. To support this claim, the author first shows that Buss' theories prove consistencies of induction-free fragments of our theories whose formulae have bounded complexity. Next, the author proves that although our theories are based on an apparently weaker logic, we can interpret theories in Buss' hierarchy by our theories using a simple translation.

1 Background

One of the central open questions in bounded arithmetic is whether Buss' hierarchy $S_2^1 \subseteq T_2^1 \subseteq S_2^2 \subseteq T_2^2 \subseteq \cdots$ of theories of bounded arithmetic collapses [5] or not. Since it is known that collapse of Buss' hierarchy implies the collapse of the polynomial-time hierarchy [7], demonstration of the non-collapse of the theories in Buss' hierarchy could be one way to establish the non-collapse of the polynomial-time hierarchy. A natural way to demonstrate non-collapse of the theories in Buss' hierarchy would be to identify one of these theories that proves (some appropriate formulation of) a statement of the consistency of some theory below it in the hierarchy.

Here, it is clear that we need a delicate notion of consistency because of several negative results that have already been established. The "plain" consistency statement cannot be used to separate the theories in Buss' hierarchy, since Paris and Wilkie [16] show that S_2 $(\equiv \bigcup S_2^i)$ cannot prove the consistency of Robinson Arithmetic Q. Apparently, this result stems more from the use of predicate logic than from the strength of the base theory. However, Pudlák [12] shows that S_2 cannot prove the consistency of proofs that are carried out within S_2^1 and are comprised entirely of bounded formulae. Even if we restrict our attention to the induction-free fragment of bounded arithmetic, we cannot prove the consistency of such proofs, as shown by Buss and Ignjatović [6]. More precisely, Buss and Ignjatović prove that S_2^i cannot prove the consistency of proofs that are comprised entirely of Σ_i^b and Π_i^b formulae and use only BASIC axioms (the axioms in Buss' hierarchy other than induction) and the rules of inference of predicate logic.

Therefore, if we want to demonstrate non-collapse of the theories in Buss' hierarchy, we should consider a weaker notion of consistency and/or a weaker theory. A number of attempts of this type have been made, both on the positive side (those that establish provability of consistency of some kind) and on the negative side (those that establish non-provability of consistency). On the positive side, Krajíček and Takeuti [8] show that $T_2^i \vdash \mathrm{RCon}(T_1^i)$, where T_1^i is obtained from T_2^i by eliminating the function symbol $\#$, and $\mathrm{RCon}(T_1^i)$ is a sentence which states that all "regular" proofs carried out within T_1^i are consistent. Takeuti [14], [15] shows that there is no "small" strictly i-normal proof w of contradiction. Here, "w is small" means that w has its exponentiation 2^w. Although Takeuti allows induction in strictly i-normal proof w, the assumption that w is small is a significant restriction to w since bounded arithmetics cannot prove existence of exponentiation. Another direction is to consider cut-free provability. Paris Wilkie [16] mentioned above proves that $I\Delta_0 + \exp$ proves the consistency of cut-free proofs of $I\Delta_0$. For weaker theories than $I\Delta_0 + \exp$, we need to relativized the consistency by some cut, then we get similar results [11], [3]. For further weaker theories, Beckmann [4] shows that S_2^1 proves the consistency of $S_2^{-\infty}$, where $S_2^{-\infty}$ is the equational theory which is formalized by recursive definitions of the standard interpretations of the function symbols of S_2. Also, it is known that S_2^i proves $\mathrm{Con}(G_i)$, that is, the consistency of quantified propositional logic G_i. On the negative side, we have the results mentioned above, that is, those of Paris and Wilkie [16], Pudlák [12], and Buss and Ignjatović [6]. In addition, there are results which extend incompleteness theorem to Herbrand notion of consistency [2],[1].

2 Our approach

In the paper [17], the author introduce the theory S_2^iE ($i = -1, 0, 1, 2 \ldots$), which for $i \geq 1$ corresponds to Buss' S_2^i, and we show that the consistency of strictly i-normal proofs that are carried out only in $S_2^{-1}E$, can be proved in S_2^{i+2}.

S_2^iE is based on the following observation: The difficulty in proving the consistency of bounded arithmetic inside S_2 stems from the fact that inside S_2 we cannot define the evaluation function which, given an assignment of natural numbers to the variables, maps the terms of S_2 to their values. For example, the values of the terms $2, 2\#2, 2\#2\#2, 2\#2\#2\#2, \ldots$ increase exponentially; therefore, we cannot define the function that maps these terms to their values, since the rate of growth of every function which is definable in S_2 is dominated by some polynomial in the length of the input [10]. With a leap of logic, we consider this fact to mean that we cannot assume the existence of values of arbitrary terms in bounded arithmetic. Therefore, we must explicitly prove the existence of values of the terms that occur in any given proof.

Based on this observation, $S_2^i E$ is formulated by using *free logic* instead of the ordinary predicate calculus. Free logic is a logic which is free from ontological assumptions about the existence of the values of terms. Existence of such objects is explicitly stated by an existential predicate rather than being implicitly assumed. See [9] for a general introduction to free logic and [13] for its application to intuitionistic logic.

Using free logic, we can force each proof carried out within $S_2^{-1}E$ to somehow "contain" the values of the terms that occur in the proof. By extracting these values from the proof, we can evaluate the terms and then determine the truth value of Σ_i^b formulae. The standard argument using a truth predicate proves the consistency of strictly i-normal proofs that are carried out only in S_2^{-1}. It is easy to see that such a consistency proof can be carried out in S_2^{i+2}.

This result improve on the positive results in the previous section in that 1) unlike T_1^i or G_i, $S_2^i E$ is based on essentially the same language as S_2^i, thereby making it possible to construct a Gödel sentence by diagonalization; 2) unlike Takeuti [14], [15], we do not assume that the Gödel number of the proofs which are proved consistent are small, that is, have exponentiations, thereby making it possible to apply the second incompleteness theorem—in particular, to derive a Gödel sentence from the consistency statement; 3) unlike the results on Herbrand and cut-free provability, $S_2^i E$ has the Cut-rule, thereby, making it easy to apply the second incompleteness theorem; and 4) unlike Beckmann [4], our system is formalized in predicate logic. On the other hand, we are still unable to show that the consistency of strictly i-normal proofs is not provable within S_2^j for some $j \leq i$. In a sense, our result is an extension of that of Beckmann [4] to predicate logic, since both results are based on the fact that the proofs contain "computations" of the terms that occur in them. In fact, if we drop the Cut-rule from $S_2^{-1}E$, the consistency of strictly i-normal proofs can be proved in S_2^1 for any i. This "collapse" occurs since, roughly speaking, the combination of the Cut-rule and universal correspond substitution rule in PV.

3 Definition of $S_2^i E$

The underlining logic of system $S_2^i E$ is a first-order predicate logic with special predicate E which signifies the existence of values. The system is formalized using sequent calculus.

The vocabulary of $S_2^i E$ is obtained from that of S_2^i by adding the unary predicate symbol E and replacing the set of function symbols of S_2 with an arbitrary but finite set $\mathcal{F}$ of function symbols which denote polynomial-time computable functions. The formulae of $S_2^i E$ are built up from atomic formulae by use of the propositional connectives $\neg, \wedge, \vee$; the bounded quantifiers $\forall x \leq t, \exists x \leq t$; and the unbounded quantifiers $\forall x, \exists x$. Implication ($\supset$) is omitted from the language, and negation ($\neg$) is applied only to equality $=$ and inequality $\leq$. These restrictions appear essential to prove consistency. If there is implication (or negation applied to arbitrary formulae) in $S_2^{-1}E$, $S_2^{-1}E$ allows induction speedup [6] [12], therefore $S_2^{-1}E$ polynomially interprets $S_2^i E$,

$i \geq 0$. This allows to prove $Ef(n)$ for any polynomial time f in $S_2^{-1}E$ by a proof whose length is bounded by some fixed polynomial of length of binary representation of n. However, this contradicts the statement of soundness (Proposition 1)

We assume that there are finitely many axioms for S_2^iE. The axioms of S_2^iE must satisfy the *boundedness conditions* defined below.

Definition 1. *A sequent* $\Gamma \to \Delta$ satisfies the boundedness conditions *if it has the following three properties, where* $\vec{a}$ *are the variables that occur free in* $\Gamma \to \Delta$.

1. *All the formulae that occur in* $\Gamma \to \Delta$ *are either in the form* $t = u$, $t \neq u$, $t \leq u$, $\neg t \leq u$, Et.
2. *Every variable in* $\vec{a}$ *occurs free in at least one formula in* Γ.
3. *There is a constant* $\alpha \in \mathbb{N}$ *such that*

$$S_2^1 \vdash \max\{\vec{t_\Delta}(\vec{a})\} \leq \alpha \cdot \max\{\vec{t_\Gamma}(\vec{a})\}, \tag{1}$$

$\vec{t_\Delta}(\vec{a})$ *are the subterms of the terms that occur in* Δ *and* $\vec{t_\Gamma}(\vec{a})$ *are the subterms of the terms that occur in* Γ *(for convenience,* $\max \emptyset$ *is defined to be 1). Since the function symbols of* S_2E *are definable in* S_2^1*, we can regard the terms in* $\vec{t_\Gamma}(\vec{a})$ *and* $\vec{t_\Delta}(\vec{a})$ *as terms of* S_2^1*, hence we can regard* $\max\{\vec{t_\Gamma}(\vec{a})\} \leq \alpha \cdot \max\{\vec{t_\Delta}(\vec{a})\}$ *as an* S_2^1 *formula.*

The consistency proof only uses this property for axioms. We refer [17] for more detail on the axioms and inference rules.

4 S_2^{i+2} consistency proof of strictly i-normal proofs

In [17], we define *i-normal formula* and *strictly i-normal proof*, and in S_2^{i+2} we prove the consistency of strictly i-normal proofs in $S_2^{-1}E$. The consistency proof is based on the facts that we can produce a Σ_i^b formula that constitutes a truth definition for i-normal formulae and we can apply the Σ_{i+2}^b- PIND rule to prove the soundness of strictly i-normal proofs in $S_2^{-1}E$. The idea is that to use a term t in an $S_2^{-1}E$ proof, we first need to prove that Et holds. To do this, we show that for a given assignment ρ of values to the variables in t, the value of t is bounded by the size of the proof of Et plus the size of ρ. Therefore, we can define a valuation function for terms and a truth definition for the formulae in the proof. Once we obtain the truth definition, consistency is easy to prove.

First, we define *bounded valuation* of terms.

Definition 2. *Let* t *be a term of* S_2^iE*, let* ρ *be an assignment for variables of* t*, and let* $u \in \mathbb{N}$*. A* ρ-valuation tree for t which is bounded by u *is a tree* w *that satisfies the following conditions.*

1. *Every node of w is of the form $\langle \ulcorner t_j \urcorner, c\rangle$ where t_j is a subterm of t, $c \in \mathbb{N}$, and $c \leq u$.*

2. *Every leaf of w is either $\langle \ulcorner 0 \urcorner, 0\rangle$ or $\langle \ulcorner a \urcorner, \rho(a)\rangle$ for some variable a in the domain of ρ.*

3. *The root of w is $\langle \ulcorner t \urcorner, c\rangle$ for some $c \leq u$.*

4. *If $\langle \ulcorner f(t_1, \dots, t_n) \urcorner, c\rangle$ is a node of w, its children are $\langle \ulcorner t_1 \urcorner, d_1\rangle, \dots, \langle \ulcorner t_n \urcorner, d_n\rangle$ which satisfy the condition $c = f(d_1, \dots, d_n)$.*

If the root of a ρ-valuation tree w for t is $\langle \ulcorner t \urcorner, c\rangle$, we say *the value of w is c.*

The statement that t converges to the value c (and $c \leq u$) is defined by the formula which expresses that the following relation (which we denote by $v(\ulcorner t \urcorner, \rho) \downarrow_u c$) holds: "$\exists w \leq s(\ulcorner t \urcorner, u)$ such that w is (the Gödel number of) a ρ-valuation tree for t which is bounded by u and has root $\langle \ulcorner t \urcorner, c\rangle$," where $s(\ulcorner t \urcorner, u)$ is a term which bounds (the Gödel numbers of) all ρ-valuation trees for t which are bounded by u. This formula appears to be Σ_1^b, since there a leading $\exists$, but actually this formula is Δ_1^b since for each t and rho, w is uniquely determined and polynomially computable.

Using the notion of ρ-valuation tree, we give a "bounded" truth definition of i-normal formulae.

First, we present a truth definition for quantifier-free formulae. Since logical symbols can be arbitrarily nested, we follow the same strategy that was used in our definition of valuation for terms. We attach a truth value to each node of a subformula tree, and we define the value attached to the root (the node that represents the entire formula) as the truth value of the formula.

Definition 3. *Let A be a quantifier-free formula of $S_2^{-1}E$, let ρ be an assignment for free variables of A, and let $u \in \mathbb{N}$. A* ρ-truth tree for A which is bounded by u *is a tree w that satisfies the following conditions.*

Every leaf of w has one of the following five forms (where in each form the possible values of ϵ are 0 and 1): $\langle \ulcorner t_1 \leq t_2 \urcorner, \epsilon\rangle$, $\langle \ulcorner t_1 \not\leq t_2 \urcorner, \epsilon\rangle$, $\langle \ulcorner t_1 = t_2 \urcorner, \epsilon\rangle$, $\langle \ulcorner t_1 \neq t_2 \urcorner, \epsilon\rangle$, $\langle \ulcorner Et \urcorner, \epsilon\rangle$.

For a leaf of the form $\langle \ulcorner t_1 \leq t_2 \urcorner, \epsilon\rangle$, $\epsilon = 1$ if $\exists c_1, c_2 \leq u$, $v(\ulcorner t_1 \urcorner, \rho) \downarrow_u c_1$, $v(\ulcorner t_2 \urcorner, \rho) \downarrow_u c_2$, and $c_1 \leq c_2$; otherwise, $\epsilon = 0$.

For a leaf of the form $\langle \ulcorner t_1 \not\leq t_2 \urcorner, \epsilon\rangle$, $\epsilon = 1$ if $\exists c_1, c_2 \leq u$, $v(\ulcorner t_1 \urcorner, \rho) \downarrow_u c_1$, $v(\ulcorner t_2 \urcorner, \rho) \downarrow_u c_2$, and $c_1 \not\leq c_2$; otherwise, $\epsilon = 0$.

The conditions that must be satisfied by a leaf of the form $\langle \ulcorner t_1 = t_2 \urcorner, \epsilon\rangle$ or $\langle \ulcorner t_1 \neq t_2 \urcorner, \epsilon\rangle$ are the obvious analogues of those for $\langle \ulcorner t_1 \leq t_2 \urcorner, \epsilon\rangle$ and $\langle \ulcorner t_1 \not\leq t_2 \urcorner, \epsilon\rangle$, respectively.

For a leaf of the form $\langle \ulcorner Et \urcorner, \epsilon\rangle$, $\epsilon = 1$ if $\exists c \leq u$, $v(\ulcorner t \urcorner, \rho) \downarrow_u c$; otherwise, $\epsilon = 0$.

Every intermediate node r of w is of the form $\langle \ulcorner A_1 \wedge A_2 \urcorner, \epsilon\rangle$ or $\langle \ulcorner A_1 \vee A_2 \urcorner, \epsilon\rangle$, where the children of r are the nodes $\langle \ulcorner A_1 \urcorner, \epsilon_1\rangle$ and $\langle \ulcorner A_2 \urcorner, \epsilon_2\rangle$.

For a node of the form $\langle \ulcorner A_1 \wedge A_2 \urcorner, \epsilon\rangle$, $\epsilon = 1$ if $\epsilon_1 = 1$ and $\epsilon_2 = 1$; otherwise, $\epsilon = 0$.

For a node of the form $\langle \ulcorner A_1 \vee A_2 \urcorner, \epsilon\rangle$, $\epsilon = 1$ if $\epsilon_1 = 1$ or $\epsilon_2 = 1$; otherwise, $\epsilon = 0$.

The root of w is $\langle \ulcorner A \urcorner, \epsilon\rangle$ for some $\epsilon \in \{0, 1\}$.

The truth of a quantifier-free formula A is defined by the Σ_1^b formula $T_{-1}(u, \ulcorner A \urcorner, \rho)$ which expresses that "$\exists w \leq s(\ulcorner A \urcorner, u)$ such that w is (the Gödel number of) a ρ-truth tree for A which is bounded by u and has root $\langle \ulcorner A \urcorner, 1 \rangle$," where $s(\ulcorner A \urcorner, u)$ is a term which bounds (the Gödel numbers of) all ρ-truth trees for A which are bounded by u.

Next, we would like to present a truth definition for Σ_i^b formulae. However, since it is technically difficult to do this for general Σ_i^b formulae, we restrict our definition to *i-normal* formulae. Since $i \in \{-1, 0, 1, 2, \ldots\}$, we have -1-normal formulae, 0-normal formulae, 1-normal formulae, 2-normal formulae, and so on.

Definition 4 (pure i-normal). *Let $i \geq -1$, and let $A(\vec{a})$ be a formula.*

If $i = -1$, $A(\vec{a})$ is pure -1-normal *if $A(\vec{a})$ is quantifier free.*

If $i \geq 0$, $A(\vec{a})$ is pure i-normal *if it is of the form*

$$\exists x_1 \leq t_1(\vec{a}) \forall x_2 \leq t_2(\vec{a}, x_1) \cdots$$
$$Q_i x_i \leq t_i(\vec{a}, x_1, \ldots, x_{i-1}) Q_{i+1} x_{i+1} \leq |t_{i+1}(\vec{a}, x_1, \ldots, x_i)| . A_0(\vec{a}, x_1, \ldots, x_{i+1}),$$

where Q_i is $\forall$ if i is even, and $\exists$ if i is odd, and $A_0(\vec{a}, x_1, \ldots, x_{i+1})$ is quantifier free and does not contain the predicate E.

Definition 5 (i-normal). *If $i = -1$, $A(\vec{a})$ is* i-normal *if it is quantifier free.*

If $i \geq 0$, $A(\vec{a})$ is i-normal *if it is a subformula of a pure i-normal formula or is Et for some term t. In other words, $A(\vec{a})$ is either an E-form, a quantifier-free formula that does not contain E, or a formula of the form*

$$Q_j x_j \leq t_j(\vec{a}, x_1, \ldots, x_{j-1}) \cdots Q_i x_i \leq t_i(\vec{a}, x_1, \ldots, x_{i-1})$$
$$Q_{i+1} x_{i+1} \leq |t_{i+1}(\vec{a}, x_1, \ldots, x_i)| . A_0(\vec{a}, x_1, \ldots, x_{i+1}), \quad (2)$$

where $A_0(\vec{a}, x_1, \ldots, x_{i+1})$ is quantifier free and does not contain E; $1 \leq j \leq i+1$; and for every k with $j \leq k \leq i+1$, Q_k is either $\forall$ or $\exists$, according as k is even or odd. If $j = i+1$, the above formula is $Q_{i+1} x_{i+1} \leq |t_{i+1}(\vec{a}, x_1, \ldots, x_i)| . A_0(\vec{a}, x_1, \ldots, x_{i+1})$.

The following is a truth definition $T_i(u, \ulcorner B \urcorner, \rho)$ for i-normal formulae B. First, we define a truth definition $T_{i,l}$ for i-normal forms with l quantifiers.

Definition 6. *Let $i \geq -1$, let B be an i-normal formula with l quantifiers. Note that $0 \leq l \leq i+1$. We define $T_{i,l}(u, \ulcorner B \urcorner, \rho)$ by recursion on l in the meta-language.*

If $l = 0$, then B is quantifier free, so $T_i(u, \ulcorner B \urcorner, \rho) \equiv T_{-1}(u, \ulcorner B \urcorner, \rho)$.

If $l \geq 1$, then

$$B \equiv Q_j x_j \leq t . A(\vec{a}, x_1, \ldots, x_j),$$

where $j = i + 2 - l$; $t \equiv t_j(\vec{a}, x_1, \ldots, x_{j-1})$ if $j < i+1$, and $t \equiv |t_{i+1}(\vec{a}, x_1, \ldots, x_i)|$ if $j = i+1$; and $A(\vec{a}, x_1 \ldots, x_j)$ is an i-normal formula with $l-1$ quantifiers. Assume *that we have defined* $T_{i,l-1}(u, \ulcorner C \urcorner, \rho)$ *for all i-normal formulae* C *with* $l-1$ *quantifiers. We define* $T_{i,l}(u, \ulcorner B \urcorner, \rho)$ *to be the following formula.*

$$\exists c \leq u, v(\ulcorner t \urcorner, \rho) \downarrow_u c \wedge Q_j d_j \leq c.T_i(u, \ulcorner A(\vec{a}, x_1, \ldots, x_j) \urcorner, \rho[x_j \mapsto d_j])$$

Then, let $\mathrm{INQ}(\ulcorner B \urcorner, l)$ *be a formula which represents "B is an i-normal form with l quantifiers". we define* $T_i(u, \ulcorner B \urcorner, \rho)$ *as*

$$\{\mathrm{INQ}(\ulcorner B \urcorner, 0) \supset T_{i,0}(u, \ulcorner B \urcorner, \rho)\} \vee \ldots \vee \{\mathrm{INQ}(\ulcorner B \urcorner, i+1) \supset T_{i,i+1}(u, \ulcorner B \urcorner, \rho).\} \quad (3)$$

Since we can contract successive $\exists$ *quantifiers into a single* $\exists$ *quantifier,* $T_i(u, \ulcorner B \urcorner, \rho)$ *is* Σ^b_{i+1}.

Finally, we prove the (sort of) the soundness of S^i_2E-proofs. However, since we restrict our attention to i-normal formulae, we can consider only *strictly i-normal proofs* for the soundness proof.

Definition 7. *An* S^{-1}_2E *proof is* strictly i-normal *if all formulae contained in the proof are i-normal. The property "w is (the Gödel number of) a strictly i-normal proof tree for* $\Gamma \to \Delta$*" is* Δ^b_1*-definable. We write* i-$\mathrm{Prf}(w, \ulcorner \Gamma \to \Delta \urcorner)$ *for the* Δ^b_1 *formula that defines this property.*

Proposition 1. *Let* Env *be the ternary relation that holds of precisely the triples* $(\rho', \ulcorner \sigma \urcorner, u)$ *where* σ *is a term, a formula, or a sequent;* ρ' *is an assignment for free variables* σ*;* $u \in \mathbb{N}$*; for every variable* x *of* S^i_2E*, there is a pair* $(\ulcorner x \urcorner, n)$ *in* ρ' *(for some* $n \in \mathbb{N}$*) if and only if* x *occurs free in* σ*; and* $\rho'(x) \leq u$ *for every variable* x *that occurs free in* σ*. We identify assignments with their Gödel numbers; therefore, we regard* Env *as a ternary relation on* $\mathbb{N}$.

Let $u \in \mathbb{N}$*, and let* σ *be a term, a formula, or a sequent.* $\mathrm{BdEnv}(\ulcorner \sigma \urcorner, u)$ *denotes the greatest* $m \in \mathbb{N}$ *which is (the Gödel number of) an assignment* ρ' *such that* $\mathrm{Env}(\rho', \ulcorner \sigma \urcorner, u)$ *holds.*

Let $\Gamma \to \Delta$ *be a sequent comprised entirely of i-normal formulae, and let* $u, w \in \mathbb{N}$ *such that* i-$\mathrm{Prf}(w, \ulcorner \Gamma \to \Delta \urcorner)$ *holds,* $w \leq u$*, and the binary representation of* u *is of the form* $11 \cdots 1$*, that is, all the bits are* 1*. Then for every node* r *of* w*, the following holds (where* ρ *denotes an environment as well as its Gödel number and* $\Gamma_r \to \Delta_r$ *denotes the conclusion of the subproof which corresponds a node* r*).*

$$\begin{aligned}\forall \rho \leq \mathrm{BdEnv}(\ulcorner \Gamma_r \to \Delta_r \urcorner, u)\Big[&\mathrm{Env}(\rho, \ulcorner \Gamma_r \to \Delta_r \urcorner, u) \supset \\ &\forall u' \leq u \ominus r\{[\forall A \in \Gamma_r,\ T_i(u', \ulcorner A \urcorner, \rho)] \supset [\exists B \in \Delta_r,\ T_i(u' \oplus r, \ulcorner B \urcorner, \rho)]\}\Big]\end{aligned} \quad (4)$$

Furthermore, this is derivable in S^{i+2}_2.

Proof. By induction on r, For the purpose of illustration, we consider the case where r is an axiom or Cut.

$$\frac{}{\Gamma(\vec{s}(\vec{a})) \rightarrow \Delta(\vec{s}(\vec{a}))}\ \mathrm{Ax}, \tag{5}$$

where $\Gamma(\vec{s}(\vec{a})) \rightarrow \Delta(\vec{s}(\vec{a}))$ is a substitution instance of an axiom.

Since there are only finite many axioms, we use case analysis on the axiom which derives this substitution instance. Assume that $\forall A \in \Gamma, T_i(u', \lceil A(\vec{s}(\vec{a}))\rceil, \rho)$. Let $\Gamma(\vec{b}) \rightarrow \Delta(\vec{b})$ be the axiom into which the substitution was made. This axiom satisfies the boundedness conditions (Definition 1). Moreover, its standard interpretation (Et is interpreted as trivially true formula.) is derivable in S_2^{i+2}.

By the first boundedness condition, all the formulae in Γ and Δ are quantifier free. Let $\vec{b} = b_1, \ldots, b_l$ and $\vec{s}(\vec{a}) = s_1(\vec{a}), \ldots, s_l(\vec{a})$, where $s_k(\vec{a})$ is the term that was substituted for the variable b_k in the application of the axiom rule ($k = 1, \ldots, l$). By the second boundedness condition, b_k occurs in Γ, so by Lemma 3.15 of [17], $\exists d_k \leq u'$ such that $v(\lceil s_k(\vec{a})\rceil, \rho) \downarrow_{u'} d_k$ ($k = 1, \ldots, l$), hence $\forall A \in \Gamma, T_i(u', \lceil A(\vec{b})\rceil, \rho[\vec{b} \mapsto \vec{d}])$.

Let $\vec{t}_\Gamma(\vec{b})$ be the subterms of the terms that occur in $\Gamma(\vec{b})$, and let $\vec{t}_\Delta(\vec{b})$ be the subterms of the terms that occur in $\Delta(\vec{b})$. Since all formulae occur in Γ and Δ are quantifier-free, $\vec{b}$ are all variables contained $\vec{t}_\Gamma(\vec{b})$ and $\Gamma(\vec{b})$. Since the function symbol of S_2^iE is definable in S_2^1, we can view the terms in $\vec{t}_\Gamma(\vec{b})$ and $\vec{t}_\Delta(\vec{b})$ as terms of S_2^1. By the third boundedness condition, the relation

$$\max\{\vec{t}_\Delta(\vec{b})\} \leq \alpha \cdot \max\{\vec{t}_\Gamma(\vec{b})\} \tag{6}$$

is provable in S_2^1.

Since $\forall A \in \Gamma, T_i(u', \lceil A(\vec{b})\rceil, \rho[\vec{b} \mapsto \vec{d}])$, we have $\max\{\vec{t}_\Gamma(\vec{d})\} \leq u'$. By Lemma 3.11 of [17] , we have that, for every A in Γ, $A(\vec{d})$ is true (in the meta-language). Since $\Gamma(\vec{d}) \rightarrow \Delta(\vec{d})$ holds (in the meta-language), there is some B in Δ such that $B(\vec{d})$ is true (in the meta-language). Since we can take α to be 4, we have $\max\{\vec{t}_\Delta(\vec{d})\} \leq 4 \cdot u' \leq u' \oplus r$. Let $\vec{c} = FV(B(\vec{b}))$. Then $T_i(u' \oplus r, \lceil B(\vec{c})\rceil, \rho[\vec{c} \mapsto \vec{d}]$ holds by Lemma 3.18 of [17]. By Lemma 3.5 of [17] and the fact that $v(\lceil s_k(\vec{a})\rceil, \rho) \downarrow_{u'} d_k$ ($k = 1, \ldots, l$), we obtain $v(\lceil s_k(\vec{a})\rceil, \rho) \downarrow_{u' \oplus r} d_k$. Using that result and Lemma 3.15 of [17], we have $T_i(u' \oplus r, \lceil B(\vec{s}(\vec{a}))\rceil, \rho)$, so we are done.

Next, we consider the Cut rule.

$$\frac{\begin{array}{cc} \vdots\, r_1 & \vdots\, r_2 \\ \Gamma \rightarrow \Delta, A & A, \Pi \rightarrow \Lambda \end{array}}{\Gamma, \Pi \rightarrow \Delta, \Lambda}\ \mathrm{Cut} \tag{7}$$

Assume that $\forall B \in \Gamma, \Pi, T_i(u', \lceil B\rceil, \rho)$. Let $\vec{a}$ be the variables that occur free in A but do not occur free in $\Gamma_r \rightarrow \Delta_r$, and let $\rho[\vec{a} \mapsto \vec{0}]$ be the environment that extends ρ and maps every variable in $\vec{a}$ to 0 (where $\rho[\vec{a} \mapsto \vec{0}] \equiv \rho$ if $\vec{a}$ is empty). Note that $\rho[\vec{a} \mapsto \vec{0}]$ is an environment for $\Gamma_{r_1} \rightarrow \Delta_{r_1}$, and that $\rho[\vec{a} \mapsto \vec{0}]$ assigns a value less than u for every variable y that occurs free in $\Gamma_{r_1} \rightarrow \Delta_{r_1}$. Let ρ_1 be the subsequence of

$\rho[\vec{a} \mapsto \vec{0}]$ such that $\mathrm{Env}(\rho_1, \ulcorner \Gamma_{r_1} \to \Delta_{r_1} \urcorner, u)$. Then we have $\forall B \in \Gamma, T_i(u', \ulcorner B \urcorner, \rho_1)$, and $\rho_1 \leq \mathrm{BdEnv}(\ulcorner \Gamma_{r_1} \to \Delta_{r_1} \urcorner, u)$. By the induction hypothesis applied to r_1 together with the fact that $u' \leq u \ominus r \leq u \ominus r_1$, either $\exists C \in \Delta, T_i(u' \oplus r_1, \ulcorner C \urcorner, \rho_1)$ or $T_i(u' \oplus r_1, \ulcorner A \urcorner, \rho_1)$.

If there is some C in Δ such that $T_i(u' \oplus r_1, \ulcorner C \urcorner, \rho_1)$, we have $T_i(u' \oplus r_1, \ulcorner C \urcorner, \rho[\vec{a} \mapsto \vec{0}])$ because ρ_1 is a subsequence of $\rho[\vec{a} \mapsto \vec{0}]$ and by Lemma 3.17 of [17]. Furthermore, $T_i(u' \oplus r_1, \ulcorner C \urcorner, \rho)$ holds by Lemma 3.17 of [17], since none of the variables in $\vec{a}$ occurs free in C. Thus $T_i(u' \oplus r, \ulcorner C \urcorner, \rho)$ by Lemma 3.16 of [17] and the fact that $u' \oplus r_1 \leq u' \oplus r$. Hence we are done, so assume otherwise.

Then we have $T_i(u' \oplus r_1, \ulcorner A \urcorner, \rho_1)$, so $T_i(u' \oplus r_1, \ulcorner A \urcorner, \rho[\vec{a} \mapsto \vec{0}])$ holds by Lemma 3.17 of [17], because ρ_1 is a subsequence of $\rho[\vec{a} \mapsto \vec{0}]$. By our assumption about Π, we have $\forall B \in \Pi, T_i(u', \ulcorner B \urcorner, \rho[\vec{a} \mapsto \vec{0}])$ by Lemma 3.17 of [17], because ρ is a subsequence of $\rho[\vec{a} \mapsto \vec{0}]$. By Lemma 3.16 of [17], we have $\forall B \in \Pi, T_i(u' \oplus r_1, \ulcorner B \urcorner, \rho[\vec{a} \mapsto \vec{0}])$.

Note that $\rho[\vec{a} \mapsto \vec{0}]$ is an environment for $\Gamma_{r_2} \to \Delta_{r_2}$, and that $\rho[\vec{a} \mapsto \vec{0}](y) \leq u$ for every variable y that occurs free in $\Gamma_{r_2} \to \Delta_{r_2}$. Let ρ_2 be the subsequence of $\rho[\vec{a} \mapsto \vec{0}]$ such that $\mathrm{Env}(\rho_2, \ulcorner \Gamma_{r_2} \to \Delta_{r_2} \urcorner, u)$. Then we have $T_i(u' \oplus r_1, \ulcorner A \urcorner, \rho_2)$ and $\forall B \in \Pi, T_i(u' \oplus r_1, \ulcorner B \urcorner, \rho_2)$; in addition, $\rho_2 \leq \mathrm{BdEnv}(\ulcorner \Gamma_{r_2} \to \Delta_{r_2} \urcorner, u)$.

Our choice of Gödel numbering, together with the fact that r_1 and r_2 are Gödel numbers of nonempty subproofs of $\Gamma_r \to \Delta_r$, ensures that $|r_1 \oplus r_2| < |r|$. Since $u' \leq u \ominus r$, we have $|u' \oplus r_1| \leq |u \ominus r \oplus r_1| < |u \ominus (r_1 \oplus r_2) \oplus r_1| = |u \ominus r_2|$, hence $u' \oplus r_1 < u \ominus r_2$.

By the induction hypothesis applied to r_2 together with the fact that $u' \oplus r_1 < u \ominus r_2$, there is some D in Λ such that $T_i(u' \oplus r_1 \oplus r_2, \ulcorner D \urcorner, \rho_2)$. Then we have $T_i(u' \oplus r_1 \oplus r_2, \ulcorner D \urcorner, \rho[\vec{a} \mapsto \vec{0}])$ by Lemma 3.17 of [17], because ρ_2 is a subsequence of $\rho[\vec{a} \mapsto \vec{0}]$. Furthermore, $T_i(u' \oplus r_1 \oplus r_2, \ulcorner D \urcorner, \rho)$ holds by Lemma 3.17 of [17], because none of the variables in $\vec{a}$ occurs free in D. Since $|r_1 \oplus r_2| < |r|$, we have $u' \oplus r_1 \oplus r_2 < u' \oplus r$, so $T_i(u' \oplus r, \ulcorner D \urcorner, \rho)$ by Lemma 3.16 of [17]. Hence we are done.

□

Theorerm 1. *Let i-Con $\equiv \forall w. \neg i$-Prf$(w, \ulcorner \to \urcorner)$, which states that there is no strictly i-normal proof of the empty sequent $\to$. Then*

$$S_2^{i+2} \vdash i\text{-Con} \tag{8}$$

Proof. We informally argue inside of S_2^{i+2}. Assume that i-Prf$(w, \ulcorner \to \urcorner)$ holds for some w. Let u be as in the statement of Proposition 1, let ρ be the empty environment, and let r be the root of w. Then we obtain $[\forall A \in \Gamma_r,\ T_i(u', \ulcorner A \urcorner, \rho)] \supset [\exists B \in \Delta_r,\ T_i(u' \oplus r, \ulcorner B \urcorner, \rho)]$. However, both Γ_r and Δ_r are empty. Therefore, we obtain $[\forall A \in \emptyset, T_i(u', \ulcorner A \urcorner, \rho)] \supset [\exists B \in \emptyset, T_i(u' \oplus r, \ulcorner B \urcorner, \rho)]$. Since there is no $A \in \emptyset$, the premise is true. But since there is no $B \in \emptyset$, the conclusion cannot be true. This is a contradiction. Therefore, the formula $\forall w. \neg i$-Prf$(w, \ulcorner \to \urcorner)$ holds. □

5 Bootstrapping Theorem of $S_2^i E$

In this section, we establish the correspondence between $S_2^i E$ and S_2^i. We show that $S_2^i E$ has essentially the same strength as S_2^i if $i \geq 1$. The theorem which establishes the correspondence is called the Bootstrapping Theorem (Theorem 2), following Buss' use of the term "bootstrapping" in [5], since we bootstrap from the restricted set of axioms of $S_2^i E$ to the full power of S_2^i.

We present a proof of the theorem in four "phases" of bootstrapping. In the first phase, we show that all the functions of $S_2 E$ are provably total. Each of the remaining phases applies to a particular class of inferences of S_2^i, and we show that all the inferences covered in each phase are *admissible in* $S_2^i E$ (if properly translated from S_2^i to $S_2^i E$), that is, that if all the premises of an inference covered in a given phase are provable in $S_2^i E$, then the conclusion of that inference is also provable in $S_2^i E$ (Definition 9). The Bootstrapping Theorem (Theorem 2) then follows from the fact that every inference of S_2^i is treated in some phase of the bootstrapping. Even the axioms are included in this, since an axiom is just a rule of inference with no premise.

5.1 Translation of theorems of S_2^i

In this subsection, we introduce a translation of S_2^i formulae to the language of $S_2^i E$ and state the Bootstrapping Theorem (Theorem 2).

Definition 8. *The formulae of S_2^i are translated into formulae of $S_2^i E$ by replacing every formula of the form $A \supset B$ with one of the form $\neg A \vee B$, and using De Morgan duality to replace every formula of the form $\neg A$ with a logically equivalent formula in which every subformula prefaced with the negation symbol "$\neg$" is of the form $t_1 = t_2$ or $t_1 \leq t_2$. We call this translation the* $*$-translation *and denote the $*$-translation of A by A^*. Formally, the $*$-translation is defined as follows.*

1. $(p(t_1, t_2))^* \equiv p(t_1, t_2)$ *if p is $=$ or $\leq$.*
2. $(\neg p(t_1, t_2))^* \equiv \neg p(t_1, t_2)$ *if p is $=$ or $\leq$.*
3. $(A \wedge B)^* \equiv A^* \wedge B^*$.
4. $(A \vee B)^* \equiv A^* \vee B^*$.
5. $(\neg A)^* \equiv (\overline{A})^*$, *where $\overline{A}$ is the De Morgan dual of A.*
6. $(A \supset B)^* \equiv (\overline{A})^* \vee B^*$.
7. $(\forall x \leq t.A)^* \equiv \forall x \leq t.A^*$ *and* $(\exists x \leq t.A)^* \equiv \exists x \leq t.A^*$.
8. $(\forall x.A)^* \equiv \forall x.A^*$ *and* $(\exists x.A)^* \equiv \exists x.A^*$.

Γ^* is the sequence of formulae which is obtained by applying $*$ to the formulae in the sequence Γ.

The sequent $\Gamma \to \Delta$ is translated to the sequent $(\Gamma \to \Delta)^* \equiv E\vec{a}, \Gamma^* \to \Delta^*$, where $\vec{a}$ are the variables that occur free in $\Gamma \to \Delta$.

The following theorem states that $S_2^i E$ proves the $*$-translations of sequents derivable in S_2^i if $i \geq 0$.

Theorerm 2 (Bootstrapping Theorem). *If $i \geq 1$ and S_2^i proves a sequent $\Gamma \to \Delta$, then $S_2^i E$ proves its $*$-translation $(\Gamma \to \Delta)^*$.*

The rest of this section is devoted to a proof of the Bootstrapping Theorem. To simplify the notation, we write $S_2^i E$ for $S_2^i E(\mathcal{F}, \mathcal{A})$

5.2 Bootstrapping Phase I: $S_2^i E$ proves totality of its functions.

In this subsection, we prove that if $i \geq 0$, all the functions of $S_2^i E$ are provably total, that is, that $S_2^i E \vdash E\vec{a} \to Ef\vec{a}$ for every function symbol $f \in \mathcal{F}$. The proof is by induction (in the meta-language) on the definition of f.

Proposition 2. *If $i \geq 0$, then for every n-ary function symbol f of $S_2^i E$, $S_2^i E$ proves*

$$E\vec{a} \to Ef\vec{a}, \tag{9}$$

where $\vec{a} \equiv a_1, \ldots, a_n$.

The reason for specifying that $i \geq 0$ is that in the proof we apply the PIND rule to Σ_0^b formulae of $S_2^i E$.

It follows from this proposition that if all the variables in a term of $S_2^i E$ converge, then the term itself converges.

Corollary 1. *Let t be a term of $S_2^i E$. If $a_1, \ldots, a_n$ are the variables that occur in t, then the following holds if $i \geq 0$.*

$$S_2^i E \vdash Ea_1, \ldots, Ea_n \to Et \tag{10}$$

Proof of Proposition 2. The proof is by induction on the definition of f. □

5.3 Bootstrapping Phase II : $S_2^i E$ proves $*$-translations of axioms of S_2^i

In Bootstrapping Phase II, we prove the $*$-translations of axioms of S_2^i in $S_2^i E$. There are two kinds of axioms: equality axioms and BASIC axioms.

Proposition 3. *The $*$-translations of the equality axioms of S_2^i are provable in $S_2^i E$.*

Proposition 4. *Assume that A is a BASIC axiom. Then $(\to A)^*$ (the $*$-translation of $\to A$) is derivable in $S_2^i E$.*

We omit the proofs of these Propositions.

5.4 Bootstrapping Phase III :$*$-translations of predicate logic are admissible in S_2^iE

In Bootstrapping Phase III, we prove that the $*$-translations of the inferences of predicate logic are admissible in S_2^iE.

Definition 9. *The inference*

$$\frac{\Gamma_1 \to \Delta_1 \quad \cdots \quad \Gamma_n \to \Delta_n}{\Gamma \to \Delta} \tag{11}$$

is admissible in S_2^iE *if* $\Gamma \to \Delta$ *is provable in* S_2^iE *whenever* $\Gamma_1 \to \Delta_1, \ldots, \Gamma_n \to \Delta_n$ *are provable in* S_2^iE.

Proposition 5. *If*

$$\frac{\Gamma_1 \to \Delta_1 \quad \cdots \quad \Gamma_n \to \Delta_n}{\Gamma \to \Delta} \tag{12}$$

is an inference of predicate logic, then the inference

$$\frac{(\Gamma_1 \to \Delta_1)^* \quad \cdots \quad (\Gamma_n \to \Delta_n)^*}{(\Gamma \to \Delta)^*} \tag{13}$$

is admissible in S_2^iE.

Proof. By case analysis.

□

5.5 Bootstrapping Phase IV : $*$-translation of Σ_i^b- PIND rule is admissible in S_2^iE

Finally, we prove admissibility of the $*$-translation of the Σ_i^b- PIND rule of S_2^i.

Lemma 1. *Assume that* $\Gamma, Ea, A(\lfloor \frac{1}{2}a \rfloor) \to A(a), \Delta$ *is provable in* S_2^iE*, where the variable* a *does not occur free in* $\Gamma \to \Delta$ *and* $A(a)$ *is a* Σ_i^b *formula. Then* $\Gamma, E\vec{a}, A(0) \to A(t), \Delta$ *is also provable in* S_2^iE*, where* $\vec{a}$ *are the variables that occur in the term* t.

Proof. Note that $\lfloor \frac{1}{2}s_0a \rfloor = \lfloor \frac{1}{2}s_1a \rfloor = a$ if Ea holds. Therefore, substituting s_0a and s_1a for a in $\Gamma, Ea, A(\lfloor \frac{1}{2}a \rfloor) \to A(a), \Delta$ and applying Cut with $Ea \to Es_0a$ and $Ea \to Es_1a$, we obtain $\Gamma, Ea, A(a) \to A(s_0a), \Delta$ and $\Gamma, Ea, A(a) \to A(s_1a), \Delta$, respectively. Combining $\Gamma, A(0) \to A(0), \Delta$ and the Σ_i^b- PIND -E rule, we have $\Gamma, Et, A(0) \to A(t), \Delta$. Since Et is derivable from $E\vec{a}$ (Corollary 1), we have $\Gamma, E\vec{a}, A(0) \to A(t), \Delta$. □

Proposition 6. *The* $*$*-translation of the* PIND *rule of* S_2^i*, that is, the inference*

$$\frac{E\vec{a}\{, Ea\}, \Gamma^*, A(\lfloor a/2 \rfloor)^* \to A(a)^*, \Delta^*}{E\vec{a}\{, E\vec{b}\}, \Gamma^*, A(0)^* \to A(t)^*, \Delta^*}, \tag{14}$$

is admissible in $S_2^i E$, where the variable a does not occur free in $\Gamma \to \Delta$, $A(a)$ is a Σ_i^b formula, $\vec{a}$ are the variables other than a that occur free in $\Gamma, A(\lfloor a/2 \rfloor) \to A(a), \Delta$, and $\vec{b}$ are the variables that occur in t but are not in $\vec{a}$.

The formula Ea (in the antecedent of $E\vec{a}\{, Ea\}, \Gamma^*, A(\lfloor a/2 \rfloor)^* \to A(a)^*, \Delta^*$) is enclosed in braces, as is $E\vec{b}$ (in the antecedent of $E\vec{a}\{, E\vec{b}\}, \Gamma^*, A(0)^* \to A(t)^*, \Delta^*$), to indicate that Ea and $E\vec{b}$ are not included in those antecedents unless the variable a occurs free in $A(a)$.

Proof. By Lemma 1.

□

References

[1] Zofia Adamowicz. Herbrand consistency and bounded arithmetic. *Fundamenta Mathematicae*, 2:1–15, 2002.

[2] Zofia Adamowicz and Paweł Zbierski. On Herbrand consistency in weak arithmetic. *Archive for Mathematical Logic*, 40(6):399–413, August 2001.

[3] Zofia Adamowicz and Konrad Zdanowski. Lower bounds for the provability of Herbrand consistency in weak arithmetics. *Fundamenta Mathematicae*, 212(3):191–216, 2011.

[4] Arnold Beckmann. Proving consistency of equational theories in bounded arithmetic. *Journal of Symbolic Logic*, 67(1):279–296, 2002.

[5] Samuel Buss. *Bounded Arithmetic.* Bibliopolis, Naples, 1986.

[6] Samuel R Buss and Aleksandar Ignjatović. Unprovability of consistency statements in fragments of bounded arithmetic. *Annals of pure and applied Logic*, 74:221–244, 1995.

[7] Jan Krajícek. Bounded arithmetic and the polynomial hierarchy. *Annals of Pure and Applied Logic*, 52(1-2):143–153, April 1991.

[8] Jan Krajíček and Gaisi Takeuti. On induction-free provability. *Annals of Mathematics and Artificial Intelligence*, 6:107–125, 1992.

[9] John Nolt. Free logic (Stanford encyclopedia of philosophy), 2010.

[10] Rohit Parikh. Existence and feasibility in arithmetic. *Journal of Symbolic Logic*, 36(3):494–508, 1971.

[11] Pavel Pudlák. Cuts, consistency statements and interpretations. *The Journal of Symbolic Logic*, 50(2):423–441, 1985.

[12] Pavel Pudlák. A note on bounded arithmetic. *Fundamenta mathematicae*, 136:85–89, 1990.

[13] Dana Scott. Identity and existence in intuitionistic logic. In Michael Fourman, Christopher Mulvey, and Dana Scott, editors, *Applications of sheaves*, volume 753 of *Lecture notes in mathematics*, pages 660–696. Springer Berlin / Heidelberg, 1979.

[14] Gaisi Takeuti. Incompleteness theorems and S_2^i versus S_2^{i+1}. In *Logic Colloquium '96: Proceedings of the Colloquium held in San Sebastián, Spain, July 9-15, 1996*, number December, pages 247–261, 1996.

[15] Gaisi Takeuti. Gödel sentences of bounded arithmetic. *The Journal of Symbolic Logic*, 2000.

[16] Alex Wilkie and Jeff Paris. On the scheme of induction for bounded arithmetic formulas. *Annals of pure and applied Logic*, 35:261–302, 1987.

[17] Yoriyuki Yamagata. Bounded Arithmetic in Free Logic. *Logical Methods in Computer Science*, 8(3), 2012.